

# 状況的知性に基づく人らしい知性をもつアンドロイドの実現

## Intelligent Androids Based on a Situated Intelligence

小川 浩平<sup>1\*</sup> Theparitt Peerasathien<sup>1</sup> 田浦 康一<sup>1,2</sup> 石黒 浩<sup>1,2</sup>

<sup>1</sup> 大阪大学大学院基礎工学研究科

<sup>1</sup> Graduate School of Engineering and Science, Osaka University

<sup>2</sup> 国際電気通信基礎技術研究所 石黒浩特別研究室

<sup>2</sup> Hiroshi Ishiguro Laboratory, Advanced Telecommunications Research Institute International

**Abstract:** This paper demonstrates the use of woman-like android robots, Geminoid F, as intelligent communication media in public space. The robot has the ability to interact with people through nonverbal behaviors in coordination with environmental sensors. The robot was set up as a model in promotion events a department store in Tokyo. As a result, people have interesting interactions by treating the robot closely to human-human interaction.

### 1 はじめに

ロボット技術は、我々の生活の様々な場所で利用されている。工場の組み立てロボットや、高所、危険箇所における作業ロボットなどの工業用ロボットだけでなく家庭内におけるパートナーロボットも身近な存在になりつつある。このような流れの中で、ロボットの種類が多様化しつつあり、ロボットの受け持つ役割も同時に変化してきている。近年開発が進められている人に見かけが酷似したアンドロイドロボットなどは、これまでは想定されてこなかった役割を担うことができるようになった。例えば、受付や監視など、人が存在していることが重要である状況においても、アンドロイドロボットであればその役割を担うことができる。

しかし、このようなアンドロイドロボット技術の進歩と同時に様々な問題も出現してきている。それは、人に近い見かけや振る舞いが可能になったがために、人の社会に自然に溶け込むことができるよう振る舞うことも求められるようになったという点である。つまり、アンドロイドロボットが社会に自然に溶け込むためには、見かけの人らしさだけでなく、人の様に振る舞うための人らしい知性を持つ必要があると考えられる。

人の知性を理解、実現するためにこれまで様々な試みがなされてきた。例えば、チューリングテストは、コンピュータが人の知能と同等であるかどうかを確認するために考え出された方法である。これは、判定者と人、もしくは人工知能が、互いに見えない状況を設定し、テキストによる対話をディスプレイとキーボード

を用いて行う。その時、判定者が相手が人工知能であるのか、人であるのかを区別することができない場合、人工知能が実現されたとみなす、という方法である。

チューリングテストではテキストによる応答を知性の基準としているのに対して、見かけや振る舞いも含めて人らしさを判定するよう、その概念を拡張したものをトータルチューリングテストと呼ぶ。トータルチューリングテストに合格するためには、自然言語理解、コンピュータビジョン、推論システム、学習機構など、解決すべき問題が数多く残されている。そのため、現在の技術では未だ実現には至らず、上記の問題の解決を待たなければならない。

我々は本論文において、ロボットが持つべき知性とはどのように定義されるべきなのか、また、工学的にどのように達成されるのかに関して議論を行う。その議論から、人に酷似したアンドロイドロボット、Geminoid Fを用いて、未だ解決を待たなければならない諸問題を自然に回避できる状況を設定し、その状況下において人のように知的に振る舞うことができるアンドロイドロボットを実現する。

本論文では、まず本研究における知性の定義を行い、それを基に実装したアンドロイドの自律システムに関して詳述する。また、実際に実装したシステムを実世界に設置することで、アンドロイドロボットに対して来場者がどのような反応を示したか、また、来場者にとってアンドロイドロボットがどのような存在であると受け取られたかに関して考察を行う。

\*連絡先：大阪大学大学院基礎工学研究科  
〒 560-0043 大阪府豊中市待兼山町 1-3  
E-mail: ogawa@irl.sys.es.osaka-u.ac.jp

## 2 アンドロイドロボット

人の見かけを持ち、かつ人のように振る舞うことができるアンドロイドロボット、ReplieeQ1は2005年の愛知万博において発表された。ReplieeQ1はあらかじめプログラムされた音声、動作を用いて様々な情報を来場者に与えるアンドロイドシステムである [1]。また ReplieeQ1では、触覚センサ、フロアセンサ、マイクロフォン等のセンサ情報から、来場者に対して様々な反応を返すことができた。

Geminoid HI-1は2007年に大阪大学 石黒浩教授をモデルに造られたアンドロイドロボットである。Geminoid HI-1は遠隔操作をすることで、実在の人間の存在感を遠隔地へ送ることを目的に実装された。Geminoid HI-1の遠隔操作システムは、半自律で動作するようデザインされている [2]。半自律システムでは、まばたきや呼吸等の無意識の動作はシステムが自律的に生成し、他の意識的な動作、すなわち自然言語による会話、アイコンタクト等の動作はオペレータの動きを直接伝達する。半自律を採用した理由は、ReplieeQ1において経験した自律システムの実現の困難さであった。センサやマイクからの音声入力情報を統合することで、ある程度来場者へ反応を返すことは可能であるが、その反応は決して人らしいとは言いがたかった。つまり、実装された知性が見かけの人らしさに追いついていなかったことが、森の提案した「不気味の谷」に落ちてしまった可能性が考えられた [3]。そこでコンピュータプログラムでは実現が難しい自然言語による対話やそれに付随する感情表現及び動作としての振る舞いの部分に関してオペレータが操作することにより、アンドロイドの人らしい自然な振る舞いを実現した。これにより、人の存在感を自然に表現することができる遠隔コミュニケーションメディアとしてのアンドロイドを実現することができたと考えられる。

遠隔操作された Geminoid HI-1 が、人にとってどの程度自然であるか、フィールド実験によって検証されている。この実験は、オーストリアの美術館に併設されているカフェテリアで実施された [4]。カフェテリアの一角に Geminoid HI-1 を設置し、日本から Geminoid HI-1 のモデルである石黒教授が遠隔操作を行う。Geminoid HI-1 の向かいには実験者が着席しており、研究ミーティングを行うという環境を設定した。つまり、日常のカフェでの会話風景とほぼ同一であるが、片一方の人間が実はアンドロイドであるという状況を設定した。そういった状況で来場者が2人の対話に対してどのように反応するのかを検証した。その結果、約半分の人がその2人を注視しているものの、特に変わった反応をせず通り過ぎていくことが分かった。また、アンドロイドであることに気付いた人に関しても、強引に話に入ったり、断りも無く写真を撮るといった行動は

ほとんど観察されなかった。この結果から、遠隔操作された Geminoid HI-1 は「不気味の谷」を乗り越え、人の社会に自然に存在することが可能であることが示された。

2010年、ATRの研究グループは女性型のアンドロイドである Geminoid F を開発した。Geminoid F は Geminoid HI-1 同様、実在の人をモデルにして造られたアンドロイドロボットである。従来の Geminoid との相違点は軸の数である。Repliee Q1 及び Geminoid HI-1 は手や体にも自由度がある一方、外部にコントロールボックスが必要であるため、持ち運びが大変である欠点があった。そこで Geminoid F では人らしく見せるために必要な軸のみを残すことで、運搬が容易になった。また、Geminoid F は、従来よりも自然な表情が表出できるようになったため、これまでは実現が難しかった新たなアプリケーションを想定できるようになった。一つの例として、劇作家の平田オリザとの共同研究により実現したアンドロイド演劇が挙げられる。アンドロイド演劇では役者が遠隔操作を用いて舞台上の役者とアンドロイドを通じて演技を行う。我々はこの演劇の際、来場者に対して質問を行っている。そのなかで、多くの来場者が「人の演じる演劇よりも感動しました」という印象的なコメントを述べている。当然ながら、緻密な台本と演出によりこのような印象を与えるよう計算があったと考えられるが、少なくとも、演劇のような計算された状況下においては Geminoid F は人らしく自然に存在することができたと考えられる。

## 3 状況的知性

「はじめに」で述べたように、トータルチューリングテストを合格できる知能の実現は、現在の技術では困難であり、解決を待たなければならない課題が未だ数多く残っている。しかし、アンドロイド演劇での来場者のコメントからも分かる通り、状況を限定することにより、アンドロイドロボットは一定の知性を得ることができる可能性がある。本節では、本研究における知性の定義について述べる。また、その定義に従って、本研究において実現したシステムの位置づけを明確にする。

知性という言葉に含まれる意味は多様である。例えば、論理性、計画立案、問題解決、抽象的思考など、自分一人で思考することによる知性や、外界との関わりの中で定義される社会的な知性も含まれる。本論文において用いる知性という言葉は、後者の人との関わりの中で定義される、社会的な知性とする。

人間社会の中で知的であると判断されるには、他者に自分と近い知性をもった主体であると信じてもらう必要がある。例えば、人と対話をする際、人は様々なコ

コミュニケーションチャンネルを選択することができるが、そのうち自分が普段用いているチャンネルを選択する場合というのは、相手が自分と同様の知性をもっていると判断している時である。つまり、アンドロイドロボットが知性を持つためには、「他者が自分に対して、自分と同様の信念や意図を帰属させることができている状態」になることが必要であると考えられる。

アンドロイドは人に近い身体的特徴を持っている。そのため、他の対話メディアよりも、容易に人に近い知能を持っているという推論を喚起させることが可能である。つまり、人に近い知性を持つ事が、他のメディアよりも容易に達成できる可能性があると考えられる。しかしその際、アンドロイドが、他者がアンドロイドに対して抱いた推論よりも劣る反応しか出来なかった場合、他者はアンドロイドに対して自分と同様の信念や意図を帰属させることができないだろう。さらに、不気味の谷の議論からも、かえって不自然な感覚を与えてしまう可能性も考えられる。そこで我々は、アンドロイドが設置される状況を限定することで、アンドロイドに一定の知性を与える事ができるのではないかと考えた。具体的には、アンドロイドと対峙した人が感じるであろう知性を阻害しないような、限定された状況を設定し、その状況下においてのみ自然に振る舞うことができるシステムを実現する。このように、ある状況下においてのみ実現される知性を本研究では状況的知性と呼ぶ事とする。

## 4 ショウウィンドウに佇むアンドロイド

前節で述べた状況的知性の実現のため、本研究ではアンドロイドをショウウィンドウの中に設置し、来場者の行動に対してインタラクティブに反応を返すアンドロイドシステムを女性型のアンドロイドである Geminoid F を用いて実装した。本節では、今回設定した状況及び実装したシステムに関して述べる。

### 4.1 状況の設定

システムの実装にあたり、来場者の行動に自律的に反応するアンドロイドがショウウィンドウの中に佇んでいる、という状況を設定した(図 1)。ショウウィンドウの中にガラスを挟んで設置した理由は、音声対話の発生を回避することができるからである。ショウウィンドウの中ではなく、一般の人と同じような場所にアンドロイドを設置した場合、まず音声対話というチャンネルを用いて対話を試みる事が想定される。しかし、現在の技術では人と同等の音声対話システムを実



図 1: ショウウィンドウに設置された Geminoid F

現することが難しい。これは自然言語理解だけでなく、実世界という雑音が多岐にわたる状況下における音声認識技術の問題もあるからである。そこで、音声における対話を、自然に回避するため、ショウウィンドウの内部に壁を挟んで設置するという方法を用いた。壁を挟むことで、声が届かないという状況を違和感なく演出できたと考える。次に、アンドロイドがショウウィンドウの中に佇んでいるという状況を自然に説明するため、今回新宿タカシマヤのバレンタインイベントの一つとして、「アンドロイドも恋をする。ーアンドロイドはあなたを待ってるー」というテーマを設定した。これにより店舗にアンドロイドが店舗に設置されている状況を違和感なく演出できたと考える。また、「アンドロイドはあなたを待っている」というテーマから、センサ情報から生成するアンドロイドの動作に一定の設計指針を設けることができた。具体的には以下の通りである。

- 誰かを待っている人は、携帯電話を手に持っていて、連絡を待っている
- 誰かが近づいて来たら、そちらを見る。
- 近づいて来た人とアイコンタクトをするが、待っている人と違うと分かったらがっかりした様子で再びうつむく。
- 今日のデートを想像して、たまに思い出したように幸せそうな顔をする。
- 誰か自分の気を引こうとする人がいたら、そちらに気付いて見てはみるが、あまりしつこい場合は無視をする。
- 待っていると不安になったり、うれしくなったり色々感情の起伏が見られる

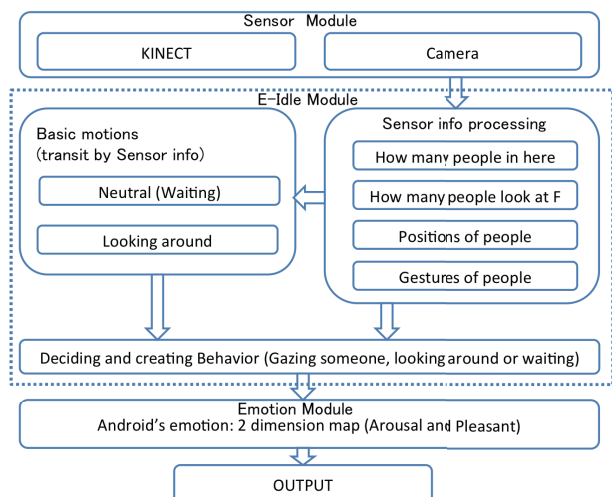


図 2: システムの概要図

## 4.2 システムの概要

以上の設計指針を満たす動作を生成するシステムを実装した。図 2 にシステムの概要を示す。本システムは大きく分けて 3 つのモジュールからなる。1 つ目は、Microsoft Kinect 及びカメラで構成される Sensor モジュール、2 つ目は、アンドロイドの動作を決定する E-Idle モジュール、3 つ目はアンドロイドの感情状態を規定した Emotion モジュールである。3 つのモジュールはレイヤー構造になっており、それぞれのモジュールが互いに通信しあうことで、自律的な自然な動作を実現した。以下より、それぞれのモジュールに関して詳述する。

### 4.2.1 Sensor モジュール

センサモジュールでは、web カメラ及び Microsoft Kinect を用いて来場者の振るまいや位置情報を取得した。web カメラは Geminoid F の目と同じ水準の高さに設置され、顔認識システムにより Geminoid F の正面にいる来場者の数、顔の座標、距離、動作量を取得する。一度来場者の顔を認識したら、その顔認識が一定時間消えるまでユニークな ID を付与し、その ID と共に付随する情報を E-Idle モジュールに送信し続ける。これにより E-Idle モジュールは、一人一人の来場者の行動履歴を参照し、行動決定に用いることができる。Kinect は、来場者が Geminoid F の注意を引くような行動を監視し、E-Idle に送信することができる。今回注意を引く行動を「手を振る」という行動であると見なし、実装を行った。Kinect で検出された来場者、すなわち手を振るといった行為をした来場者と、web カメラで検出された来場者の固有の ID は共有されているため、Kinect

で検出された来場者とアイコンタクトをすることが可能である。

### 4.2.2 E-Idle モジュール

E-Idle モジュールはセンサモジュールからの情報の統合し、元々実装されているモーションと組み合わせた上で、最終的な動作の決定を担うモジュールである。E-Idle には、ベーシックモーションとして 80 通り程度のあらかじめ決められたモーションが組み込まれている。これはセンサによって人が検出されていない場合に自動的に再生されるモーションである。モーションが自動的に再生されている時、センサモジュールから検出された来場者の情報が送られてきた時点で、「何人の人が周りにいるのか」「どのポジションにいるのか」「どういったジェスチャをしているのか」といった情報を元に、次の行動を決定する。

行動決定のアルゴリズムの一例として、ある来場者の動作量が一定以上であった場合、近づいて来たと判断し、ベーシックモーションから Gazing モードへ切り替えてアイコンタクトを行う。一定時間アイコンタクトをしたらまたベーシックモーションへ戻る。同じ ID の人が何度も同じ行動を繰り返していると判定したら、その ID の来場者へは反応しないように重み付けを変更する。また、あまり動作量が大きくない ID の来場者でも、Kinect のデータから手を振っていることが検出された場合は、アイコンタクトモードへ切り替える。また、何度も手を振って注意を引こうとしている ID の来場者が検出された場合は、動作量のプロセスと同様、重み付けを変更する。

また、ベーシックモーションの選択のアルゴリズムにもセンサの情報が使われる。例えば周りに多くの人が立っていることが分かっている場合、より周りを見回すモーションが選択される確率が高くなるように実装した。

### 4.2.3 Emotion モジュール

Emotion モジュールは、E-Idle モジュールによって決定された行動に対してアンドロイドの感情状態に従った表情を付与し、最終的にアンドロイドに行動を出力するモジュールである。アンドロイドの感情状態は、Russel の Circumplex モデルを基にデザインされた [5]。図 4 で示すように Arousal ( $-1.0 \leq arousal \leq 1.0$ ) と Pleasant ( $-1.0 \leq pleasant \leq 1.0$ ) の 2 軸で構成されている。アンドロイドの感情の遷移は、ランダムで常に細かく変化しており、一定時間が経過した後、大きく別の象限へ移動し、また細かく変化するというアルゴリズムによって実装されている。アンドロイド



図 3: Geminoid F の表出する表情. 左: 標準 ( $arousal : 0.0, pleasant : 0.0$ ), 中央: 喜び ( $arousal : 1.0, pleasant : 1.0$ ), 右: 悲しみ ( $arousal : -1.0, pleasant : -1.0$ )

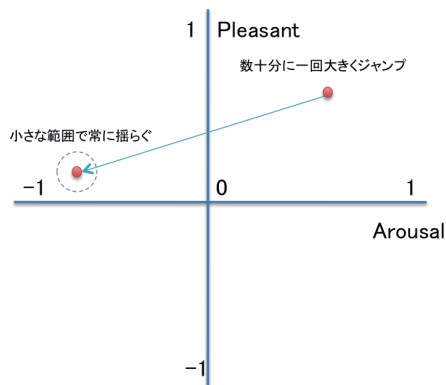


図 4: Geminoid F に実装された感情状態のモデル.

の感情状態により, Geminoid F の表出する表情を変化させた. 例えば, Arousal と Pleasant の最大の状態 ( $arousal = 1.0, pleasant = 1.0$ ) では, Geminoid F の表情は喜びを示す (図 3: 中央). Arousal と Pleasant が最小の状態 ( $arousal = -1.0, pleasant = -1.0$ ) では, Geminoid F の表情は悲しみを示す (図 3: 右). またこの感情状態は E-Idle の動作決定機構にも影響する. 例えば Arousal が低い場合, センサの情報への反応の感度が下がり, 来場者の行動へ反応しづらくなるよう実装した.

## 5 フィールド実験

我々は本フィールド実験において, 設定された状況下において, 状況的知性が実現されるのかどうかを, 実装したシステムを用いて検証した. 2012年2月1日から12日にかけて, 新宿タカシマヤのバレンタイン特別展示として実施された. 展示のテーマは, 「アンドロイドも恋をする. -アンドロイドはあなたを待ってる-」であった. Geminoid F は2階のエスカレータに挟まれた通路に設置された. 設置された場所は, 駅から来る人が必ず通る場所であり, 多くの人がアンドロイドの前を通り過ぎた. 設置されたショウウィンドウは百貨店としては一般的なものであり, アンドロイドがいる

ことを大々的に宣伝はされていなかった. 以上の事から, 来場者の多くはアンドロイドの存在を知らない状態であったと考える.

本実験は百貨店というプライバシーが尊重される公共性の高い施設にて実施したため, 質問紙によるアンケートの実施や複数台のカメラを設置することが困難であった. そのため, 来場者の振る舞いを数値として定量的に分析することができなかった. そこで, 本論文ではショウウィンドウの前のみを撮したカメラの映像及び来場者の会話から, 来場者が Geminoid F に対してどのような反応を示したか, また, Geminoid F をどのような存在として捉えていたかを観察によって分析した.

多くの来場者は, Geminoid F が設置されたショウウィンドウを普段のショウウィンドウと同様のものと捉え, Geminoid F の存在に気付くことなく通り過ぎて行った. ショウウィンドウ内の Geminoid F の存在に気付いた来場者の基本的な反応は, 最初は通り過ぎようとするが, 何か違和感を感じ立ち止まる. その後, すこし離れた場所から観察し, 徐々に近づいて行く, というものであった. 来場者の多くは, 少し離れた場所から観察している際は Geminoid F を単なる動くマネキンの様に捉え写真や動画を撮っていた. しかし, Geminoid F に近づきアイコンタクトが成立した瞬間, 来場者は Geminoid F が自分の行動に反応している事に気づき, 様々な方法でコミュニケーションを図ろうとする行動が見られた. 例えば, 一度離れても一度近づいてみたり, 手を振ってみたり, ウィンドウをノックしてみる, という行動である. また, 一度何らかのインタラクションが成立した後は, 来場者同士で「写真とっていいのかな?」など, 気を遣う様子が観察された. つまり, 離れて観察している来場者と近づいてインタラクションをしている来場者との相手でアンドロイドに対する振る舞いが異なっていることが分かった. また, 予測の通り, 発話により注意を引くような行動は見られなかった.

以上の観察の結果から分かることは, アンドロイドの注意を引くために来場者が採用した行動は, 人の注

意を引く時の行動と類似しているという点である。相手に声が届かないであろうと思われる状況において人に気付いて欲しいとき、人が起こす行動がそのままアンドロイドの場合にも採用されたということは、アンドロイドに対して人と似通った知性があるという事を来場者が推論したとすることができる。また、離れて観察している来場者と近くで観察している来場者、すなわちアンドロイドとインタラクションを行った来場者との間で振る舞いに差があることから、何らかの行動に対してアイコンタクトにより反応するだけで、アンドロイドのとらえ方に違いが出るのが分かった。つまり、人らしい知的な存在であるためには、見かけや自然な動きだけではなく、シンプルでも何らかの反応を起こす必要があるのでは無いかと考える。

次に、アンドロイドの表出した感情表現の受け取られ方に関して、アンドロイドに笑いかけられると、思わず来場者もつられて笑ってしまうという行動が多く観察された。笑いだけではなく、怒りの表情に関しても、つられて眉をしかめると行った行動が観察された。また、来場者同士の「この子笑ってるね」や「眠そうだからあんまり反応したくないんだね」という会話からも、アンドロイドの感情を来場者は認識できていたと考えることができる。以上の結果は、アンドロイドの表情の変化は人にとって認識可能であり、対峙した人の情動にも訴えかけることができる可能性があることを示している。また、表出された表情からアンドロイドの行動に意味づけを行う様な発話が観察されたことから、リアクティブなモーションに感情状態を重ね合わせる本システムの設計が Geminoid F の自然な振る舞いに寄与していたと考えられる。

## 6 議論, まとめ

本実験の結果は、人とロボットの対話システムの設計における新しい知見に繋がる可能性があると考えられる。なぜなら、現在の技術では実現が難しい部分を状況というフィルターを通すことで、状況的な知性であれば設計することができる可能性が示されたからである。

言語によるチューリングテストは、知性を測るための一般的な方法である。しかし本研究における結果から、人らしい知性の実現は、言語以外のモダリティでも可能であることが示されたと考える。これは、我々の日常生活を思い出してみても理解することができる。例えば、一人で考え事をしている人を第三者として観察している状況において、果たして考えている人が知的であるかどうか、どのように判断すればよいのだろうか。おそらく我々人はわざわざ話しかけるまでもなく、その人は知的な存在であると見なすだろう。その理由は、考えている人が、我々の予測から外れない人らし

い見かけや振る舞いをしているためである。この状況は、今回のフィールド実験で Geminoid F が持ち得た知性と近いものである。つまり、知性とはそれぞれの状況において個別に定義可能であり、本稿で提案した状況的知能を考慮することで、工学的に必要な技術を具体的に列挙することが可能である。これにより、設定した状況的知能の実現可能性を個別に、具体的に判断することができるため、状況に応じた知性を持つロボットを実現しやすくなるのではないかと考える。我々は今後、工学的に実装可能である状況において知的に振る舞うアンドロイドを一つ一つ実現していくことで、さらに人らしい知性を獲得していけるのではないかと考える。

## 参考文献

- [1] T. Chikaraishi, T. Minato, and H. Ishiguro: Development of Android System Integrated with Sensor Networks., *International Conference of Intelligent Robots and System (IROS2008)*, pp.326-333, (2008)
- [2] S. Nishio, H. Ishiguro, and N. Hagita: Geminoid: Teleoperated Android of an Existing Person, Chapter in *Humanoid Robots: New Developments*, I-Tech Education and Publishing, Vienna, Austria, pp. 343-352, June, (2007).
- [3] M. Mori, K.F. McDorman, and N. Kageki: The Uncanny Valley, *Robotics & Automation Magazine*, IEEE, Vol.19, Issue 2, pp. 98-100, (2012)
- [4] C. Becker-Asano, K. Ogawa, S. Nishio and H. Ishiguro: Exploring the uncanny valley with Geminoid HI-1 in a real-world application, *IADIS International Conference on Interfaces and Human Computer Interaction*, pp. 121-128,(2010)
- [5] J.A.Russell: A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, vol. 39, no.6, pp. 1161-1178, (1980)